



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 15, 2018 – 10:46 PM EST

PDB ID : 3IKM
Title : Crystal structure of human mitochondrial DNA polymerase holoenzyme
Authors : Lee, Y.-S.; Kennedy, W.D.; Yin, Y.W.
Deposited on : 2009-08-06
Resolution : 3.24 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030736
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030736

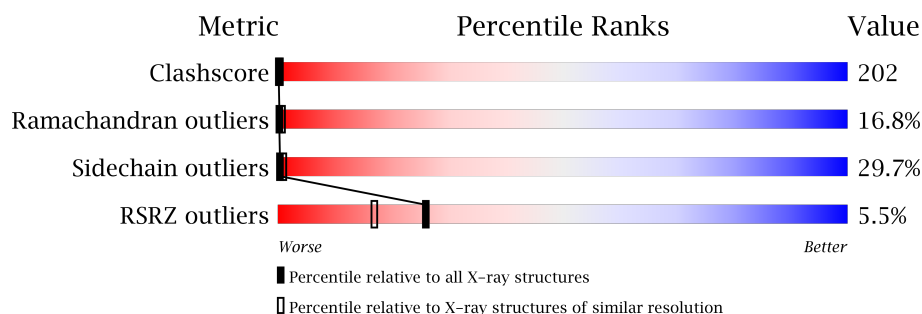
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.24 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	1383 (3.28-3.20)
Ramachandran outliers	110173	1358 (3.28-3.20)
Sidechain outliers	110143	1357 (3.28-3.20)
RSRZ outliers	101464	1252 (3.28-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1172	<div> <div>5%</div> <div>5% 45% 34% 9% 7%</div> </div>
1	D	1172	<div> <div>4%</div> <div>5% 43% 34% 11% 7%</div> </div>
2	B	427	<div> <div>7%</div> <div>11% 50% 17% 5% 17%</div> </div>
2	C	427	<div> <div>4%</div> <div>6% 54% 25% 8% 7%</div> </div>
2	E	427	<div> <div>7%</div> <div>6% 55% 19% • 17%</div> </div>
2	F	427	<div> <div>5%</div> <div>5% 54% 25% 8% 7%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 29480 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA polymerase subunit gamma-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1093	Total	C	N	O	S	0	0	0
			8681	5505	1537	1589	50			
1	D	1094	Total	C	N	O	S	0	0	0
			8695	5515	1540	1590	50			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	354	ASN	-	SEE REMARK 999	UNP P54098
A	355	SER	-	SEE REMARK 999	UNP P54098
D	354	ASN	-	SEE REMARK 999	UNP P54098
D	354A	SER	-	SEE REMARK 999	UNP P54098

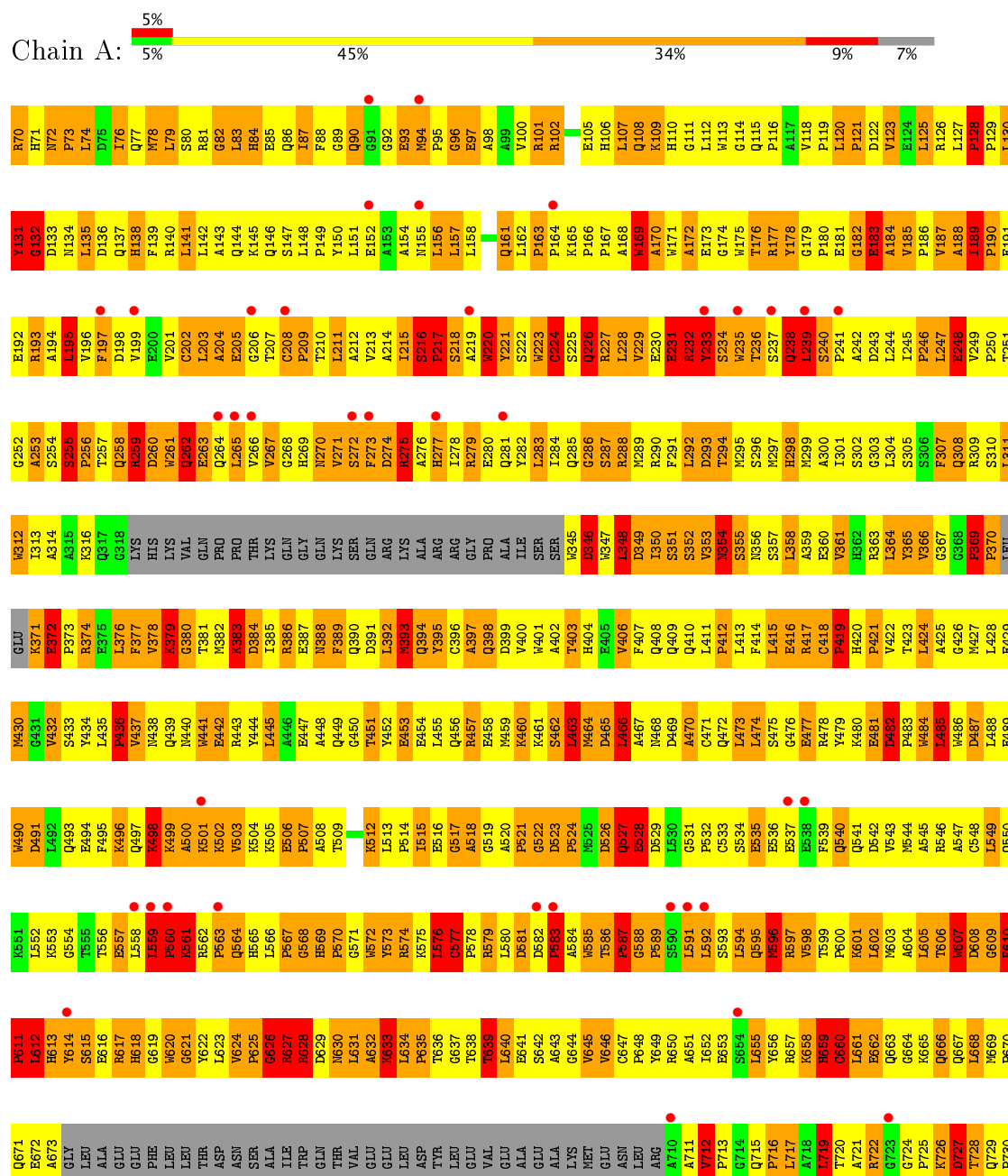
- Molecule 2 is a protein called DNA polymerase subunit gamma-2.

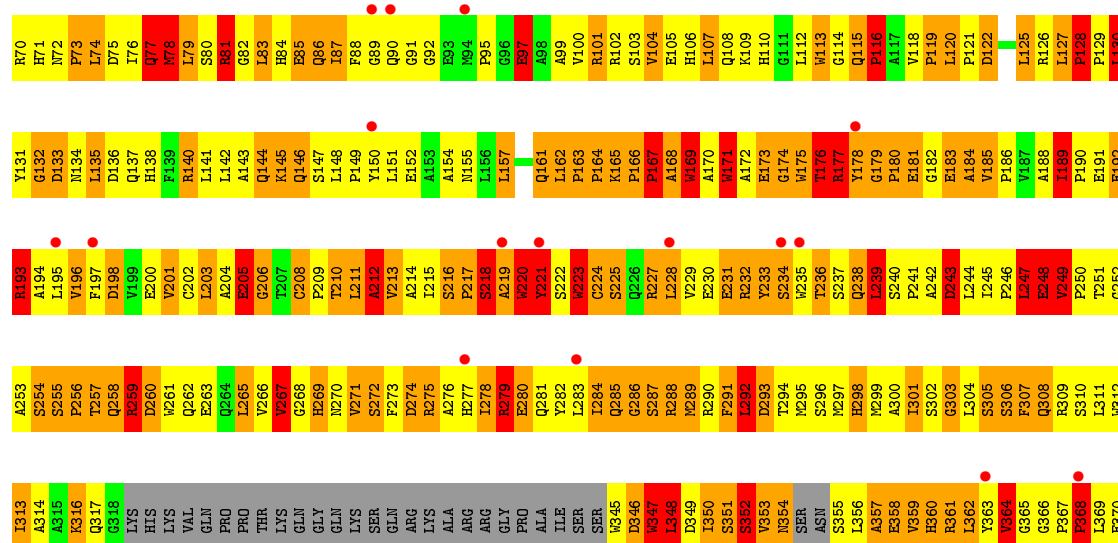
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	355	Total	C	N	O	S	0	0	0
			2871	1840	504	511	16			
2	C	396	Total	C	N	O	S	0	0	0
			3181	2031	563	571	16			
2	E	355	Total	C	N	O	S	0	0	0
			2871	1840	504	511	16			
2	F	396	Total	C	N	O	S	0	0	0
			3181	2031	563	571	16			

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

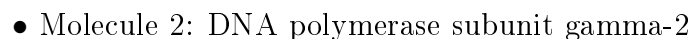
- Molecule 1: DNA polymerase subunit gamma-1





Q1214	Q1154	T1094	R1034	T974	T914	R853	E793	H733	A673	H613	K553	D491	G431	K371
G1215	I1155	S1095	K1035	Q875	A915	A854	I794	H734	GLY	Y614	G554	D492	V432	E372
E1216	T1156	R1096	S1036	Q876	F916	V855	K796	G735	LEU	S615	T595	Q493	S433	E373
A1217	N1157	V1097	Q1037	R977	G917	E856	K796	N736	ALA	R616	T556	E494	E434	R374
L1218	L1158	N1098	K1038	A978	N918	R857	I798	G737	GLU	R617	E557	F495	L435	E375
D1219	L1159	V1099	K1039	A979	N919	T858	S799	P738	GLU	H618	L558	K496	P436	L376
T1220	T1160	V1100	K1040	R980	T920	R859	S799	P739	PHI	Q497	L559	Q497	V437	F377
Y1221	T1161	V1101	W1041	R981	L921	L860	F800	N740	LEU	W620	P560	K498	W438	V378
Q1222	C1162	E1042	E1042	A982	G922	S863	W801	D741	LEU	W621	K499	K499	Q439	R379
I1223	S1103	S1043	V1043	Q983	G923	N864	R802	V742	THR	G620	R500	A500	N440	G380
I1224	F1164	S1104	V1044	Q984	G924	A865	N803	D743	ASP	L623	P563	K501	W441	T381
E1225	A1105	A1045	A1045	N985	G925	R866	H805	I744	ASN	A804	D564	K502	E442	K382
Y1226	V1106	N986	E1046	N986	S926	R867	K806	P745	SER	P625	R565	K503	R443	K383
T1227	K1167	A987	R1047	A987	A927	P867	K806	G746	ALA	G626	L566	K504	Y444	D384
G1228	L1168	A1048	A1048	A988	G928	D868	R807	T747	ILE	R627	P567	K505	L445	T385
E1229	G1169	T989	R869	R989	T929	R869	R808	W748	TRP	R628	G568	E506	A446	R386
S1230	K1170	K1050	K1050	K990	D930	V870	S809	F749	GLN	D629	H569	GLN	A447	E387
L1231	G1051	L1111	G1051	G991	L931	G871	S810	P750	THR	N630	P570	T509	A448	N388
D1232	G1052	L1112	G1052	L992	H932	S872	Q811	K751	VAL	L631	G571	A510	Q449	F389
L1233	T1053	R993	E873	R993	S933	E873	W813	L752	GLU	R633	N572	S511	Q450	Q390
R1234	V1114	N994	L874	R994	G934	K375	W814	H754	LEU	L634	K574	K512	T451	D391
E1235	A1115	R995	A936	R995	T935	A876	W815	K755	ASP	P635	K575	L513	T452	L392
S1236	M1116	R996	A936	R996	G936	A877	L816	D756	TYR	T636	L576	I515	E453	K393
V1237	K1117	N997	T937	N997	T937	W877	P817	G757	LEU	G637	C577	E516	E454	D394
G1238	W1118	S998	T938	S998	V938	W878	R818	W758	GLU	T638	P578	G517	L455	K395
F1239	L1119	D999	Q879	D999	V939	Q879	R818	W758	GLU	T638	P578	G517	Q456	C396
F1240	K1180	K1060	A980	A980	G940	A980	S819	S759	VAL	T639	R579	A518	R457	A397
S1241	E1121	L1061	R981	R981	L941	R981	A820	C760	GLU	L640	L580	G519	E458	K398
A1242	E1062	E1062	R982	R982	S942	R982	L821	N761	ALA	A520	D581	P521	M459	D399
L1243	S1063	K1063	G883	G883	R943	G883	P822	V762	GLU	S642	D582	P521	K460	V400
I1244	A1064	L1064	V884	V884	E944	V884	R823	G763	ALA	A643	P583	G522	K461	H401
L1245	V1005	H945	T885	T885	H945	T885	A824	S764	LVS	G644	A584	D523	S462	H402
L1246	R1006	A946	W825	W825	A946	W825	R826	P766	MET	V645	W585	P524	L463	T403
G1247	E1007	R947	R827	R827	R947	R827	I826	F766	GLU	V646	T586	M526	M464	H404
L1248	N1009	F949	A889	A889	F949	A889	R827	A767	ASN	C647	P587	D526	D465	E405
C1249	P1070	D890	K768	K768	D890	K768	P829	D769	LEU	P649	G588	Q527	L466	V406
L1250	P1011	V951	W881	W881	N950	D890	D830	F770	A710	R650	P589	D528	A467	F407
S1251	D1012	G952	D892	D892	G952	D892	I831	L771	A711	A651	L591	D529	W468	Q408
D1013	D1013	R953	S893	S893	R953	S893	D832	P772	V712	L652	L592	G530	D469	Q409
R1014	R1014	I954	Q894	Q894	I954	Q894	E833	K773	P713	S653	S593	P532	C471	Q410
T1015	T1015	G955	E895	E895	G955	E895	E834	N774	G714	S654	L594	P532	Q472	P412
E1016	E1016	A956	L896	L896	A956	L896	G835	E775	Q715	L655	D595	G533	L473	L413
G1017	G1017	A957	R897	R897	A957	R897	I836	D776	P716	Y656	N596	D542	L474	L414
C1018	C1018	G958	I898	I898	G958	I898	K837	G777	L717	R657	R597	E537	S475	L415
W1019	W1019	Q959	A899	A899	Q959	A899	G838	T778	A718	K658	V598	E538	G476	E416
I1020	I1020	P960	A900	A900	P960	A900	R839	L779	L719	H659	T599	F539	E477	R417
S1021	S1021	F961	V901	V901	F961	V901	I840	Q780	T720	C660	P600	Q540	R478	C418
A1022	A1022	A962	L902	L902	A962	L902	L841	A781	A721	L661	K601	Q541	Y479	P419
Q1023	Q1023	E963	G903	G903	E963	G903	P842	G782	R722	E662	L602	D542	K480	H420
D1024	D1024	R964	D904	D904	R964	D904	Q843	P783	G723	Q663	H603	V543	E481	P421
L1025	L1025	I965	A905	A905	I965	A905	K843	G784	G724	G664	A604	V544	D482	V422
R1026	R1026	L966	H906	H906	L966	H906	W845	E785	P725	K665	L605	A545	P483	T423
A1027	A1027	N967	F907	F907	N967	F907	T846	A786	K726	Q666	T606	R546	W484	L424
V1028	V1028	Q968	A908	A908	Q968	A908	A847	S787	D728	Q667	D607	A547	L485	A425
Q1029	Q1029	P969	G909	G909	P969	G909	G848	T788	T728	L668	D608	C548	W486	A426
R1030	R1030	N970	H910	H910	N970	H910	T849	P789	Q729	M669	G609	L549	D487	M427
E1031	E1031	H971	R912	R912	H971	R912	I850	R790	P730	P670	F611	Q550	L488	L428
T1032	T1032	R972	G912	G912	R972	G912	T851	A791	S731	Q671	P611	Q551	E489	E429
A1033	A1033	L973	C913	C913	L973	C913	R852	L792	Y732	E672	L612	L552	W490	M430

- Molecule 2: DNA polymerase subunit gamma-2







4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	138.39Å 138.39Å 226.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.92 – 3.24 46.92 – 3.25	Depositor EDS
% Data completeness (in resolution range)	89.5 (46.92-3.24) 89.5 (46.92-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.12 (at 3.25Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.284 , 0.303 0.362 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	93.9	Xtriage
Anisotropy	0.065	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 64.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	0.008 for -h,-k,l 0.459 for h,-h-k,-l 0.009 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	29480	wwPDB-VP
Average B, all atoms (Å ²)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.12% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.72	10/8913 (0.1%)	1.33	149/12100 (1.2%)
1	D	0.81	18/8927 (0.2%)	1.45	161/12118 (1.3%)
2	B	0.60	2/2944 (0.1%)	0.99	16/3981 (0.4%)
2	C	0.68	0/3262	1.31	47/4411 (1.1%)
2	E	0.67	6/2944 (0.2%)	1.11	23/3981 (0.6%)
2	F	0.64	0/3262	1.28	37/4411 (0.8%)
All	All	0.72	36/30252 (0.1%)	1.31	433/41002 (1.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	D	1	3
2	B	0	1
2	C	0	1
2	E	1	2
2	F	0	1
All	All	2	12

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	358	GLU	CG-CD	12.87	1.71	1.51
2	E	385	LEU	CA-CB	-11.49	1.27	1.53
1	D	352	SER	N-CA	9.45	1.65	1.46
1	A	232	ARG	N-CA	-8.85	1.28	1.46
2	E	382	LYS	N-CA	-8.69	1.28	1.46

The worst 5 of 433 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	382	LYS	N-CA-C	16.56	155.72	111.00
2	E	385	LEU	N-CA-C	16.34	155.11	111.00
1	D	248	GLU	CA-C-N	-14.80	84.64	117.20
2	B	280	ASP	C-N-CA	14.32	157.51	121.70
1	A	1212	ILE	C-N-CD	-12.99	92.03	120.60

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	249	VAL	CA
2	E	385	LEU	CA

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1179	PHE	Sidechain
1	A	221	TYR	Sidechain
1	A	239	LEU	Mainchain
1	A	395	TYR	Sidechain
2	B	417	TYR	Sidechain

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	8681	0	8521	3592	8
1	D	8695	0	8540	3900	7
2	B	2871	0	2862	809	0
2	C	3181	0	3168	1419	1
2	E	2871	0	2862	945	0
2	F	3181	0	3168	1467	4
All	All	29480	0	29121	11814	10

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 202.

The worst 5 of 11814 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:352:SER:C	1:D:352:SER:CA	1.74	1.55
1:D:460:LYS:N	1:D:460:LYS:HE2	1.20	1.48
2:E:432:TYR:HB2	2:E:437:ILE:CD1	1.51	1.39
1:D:352:SER:CA	1:D:358:GLU:OE2	1.73	1.36
1:D:914:THR:O	1:D:918:TRP:HB3	1.24	1.35

The worst 5 of 10 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:673:ALA:O	2:F:318:ASN:O[2_734]	1.91	0.29
1:A:673:ALA:C	2:F:318:ASN:O[2_734]	1.98	0.22
1:A:1034:ARG:NE	2:F:450:ASN:ND2[1_545]	2.04	0.16
1:A:994:TRP:NE1	1:D:522:GLY:N[1_545]	2.06	0.14
1:D:1232:GLU:OE2	2:F:223:GLN:OE1[2_844]	2.11	0.09

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1085/1172 (93%)	629 (58%)	254 (23%)	202 (19%)	0	0
1	D	1086/1172 (93%)	645 (59%)	234 (22%)	207 (19%)	0	0
2	B	349/427 (82%)	205 (59%)	94 (27%)	50 (14%)	0	1
2	C	394/427 (92%)	254 (64%)	84 (21%)	56 (14%)	0	1
2	E	349/427 (82%)	219 (63%)	79 (23%)	51 (15%)	0	1
2	F	394/427 (92%)	251 (64%)	93 (24%)	50 (13%)	0	2
All	All	3657/4052 (90%)	2203 (60%)	838 (23%)	616 (17%)	0	1

5 of 616 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	ASN
1	A	76	ILE
1	A	96	GLY
1	A	163	PRO
1	A	169	TRP

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	918/988 (93%)	611 (67%)	307 (33%)	0	0
1	D	919/988 (93%)	584 (64%)	335 (36%)	0	0
2	B	317/380 (83%)	259 (82%)	58 (18%)	2	9
2	C	350/380 (92%)	260 (74%)	90 (26%)	0	2
2	E	317/380 (83%)	270 (85%)	47 (15%)	3	16
2	F	350/380 (92%)	245 (70%)	105 (30%)	0	1
All	All	3171/3496 (91%)	2229 (70%)	942 (30%)	0	1

5 of 942 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	369	ARG
1	D	308	GLN
2	F	231	ILE
2	C	407	LEU
1	D	145	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 119 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	454	HIS
1	D	394	GLN
2	F	84	GLN
1	D	86	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	155	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	1093/1172 (93%)	0.20	58 (5%) 27 17	27, 80, 112, 124	0
1	D	1094/1172 (93%)	0.12	47 (4%) 36 25	27, 81, 113, 124	0
2	B	355/427 (83%)	0.40	29 (8%) 12 9	81, 111, 129, 141	0
2	C	396/427 (92%)	0.28	19 (4%) 31 21	42, 85, 113, 127	1 (0%)
2	E	355/427 (83%)	0.38	30 (8%) 11 8	64, 108, 129, 144	0
2	F	396/427 (92%)	0.17	21 (5%) 27 17	42, 87, 114, 128	0
All	All	3689/4052 (91%)	0.22	204 (5%) 26 17	27, 88, 120, 144	1 (0%)

The worst 5 of 204 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	267	ALA	7.4
1	D	583	PRO	6.8
2	F	327	GLY	6.7
1	D	277	HIS	6.1
1	A	563	PRO	6.1

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.